National College of Ireland

Enhancing Ethereum Fraud Detection Accuracy with Sparse-Attention-Based Model

MSc Research Project

MSc Data Analytics-C

Vineeth Kumar Reddy Chandravathi

Student ID: x23199091

School of Data Analytics

National College of Ireland

Supervisor: Harshani Nagahamulla

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name:	Vineeth Kumar Reddy Chandravathi
Student ID:	X23199091
Programm:	MSc in Data Analytics **Year** : 2024-2025
Module:	MSc Research Project
Supervisor:	Harshani Nagahamulla
Submission Due Date:	15/09/2025
Project Title:	Enhancing Ethereum Fraud Detection Accuracy with Sparse-Attention-Based-Model
Word Count:	7208
Page Count:	21

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: Vineeth Kumar Reddy Chandravathi

Date: 15/09/2025

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	✓
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	✓
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	✓

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Enhancing Ethereum Fraud Detection Accuracy with Sparse-Attention-Based Model

Vineeth Kumar Reddy Chandravathi

Student ID: x2319909

Abstract

As decentralized finance (DeFi) expands, Ethereum's role as the backbone for digital asset exchange, smart contracts, and financial protocols has grown—but so has its exposure to fraud. Phishing, money laundering, and malicious contracts exploit its openness. Existing ML and deep learning models often lack the balance between speed, accuracy, and explainability needed for real-time blockchain analysis. High-performing models like transformers are accurate but too resource-heavy and opaque. This study leverages TabNet—a sparse-attention deep learning model optimized for tabular Ethereum transaction data. It dynamically selects relevant features during training, enhancing both interpretability and efficiency. With an accuracy of 0.86, precision of 0.80, and F1-score of 0.79, TabNet outperforms traditional models in fraud detection while remaining lightweight and transparent. Its feature-level insights make it ideal for environments where trust, latency, and transparency are crucial. The results position TabNet as a scalable, practical alternative for fraud detection in blockchain ecosystems.

1 Introduction

1.1 Background and Problem Context

The blistering growth of blockchain technologies and decentralized finance (DeFi) platforms has reshaped the financial systems across the globe by making the trustless peer-to-peer transactions possible. Ethereum is one such platform which has emerged as a hub because of its programmability and extensive use in smart contracts. Nevertheless, malicious actors have also been interested in this innovation, which is why fraudulent schemes like Ponzi schemes, phishing attacks, and fraudulent smart contracts have become increasingly common (Onu et al., 2023; Choi & Buu, 2024). The pseudonymous, immutable and distributed nature of Ethereum renders traditional fraud monitoring tools ineffective and there is a need to introduce intelligent, adaptive fraud detection mechanisms

.

The rule-based systems are usually fixed and cannot be changed to the changing patterns of threats. Machine learning (ML) and deep learning (DL) approaches have become an alternative as fraudsters come up with more advanced solutions because they can learn and identify slight anomalies (Ehsan et al., 2024). Most of these methods are not suitable to run in real-time, large-

scale Ethereum networks due to high computational costs, long training times, or a lack of interpretability barriers (Umer et al., 2023).

Several recent studies have employed advanced transformer architectures and graph-based models to improve detection accuracy. For instance, transformer-based systems excel in capturing sequential and contextual dependencies in transaction data (Olusegun & Yang, 2024; Nam et al., 2025), while Graph Neural Networks (GNNs) model the relational structure of blockchain networks with strong predictive performance (Kanezashi et al., 2022). Nevertheless, such models tend to be computationally expensive and need voluminous labeled data, which limits their applicability in time-sensitive or resource-limited applications.

In addition, most of the current models are either highly accurate or highly interpretable but not both. Ensemble methods can be stacked to further enhance performance, but they are black boxes, and it cannot provide much information about the contribution of features (Md et al., 2023). This indicates a major research gap: the lack of a lightweight, explainable, and performance-efficient model that is optimized to tabular transaction data, the most common structure of most Ethereum fraud detection datasets.

1.2 Motivation

With the exponential growth of decentralized finance (DeFi) platforms, Ethereum has become a major hub for digital asset transactions. However, this rise in popularity has also led to a surge in fraudulent activities, including phishing scams, money laundering, and malicious smart contracts. Detecting such fraud is inherently complex due to the highly dynamic, high-dimensional, and imbalanced nature of transaction data. Existing machine learning and deep learning models either fail to generalize across evolving fraud patterns or lack the interpretability and computational efficiency needed for real-time applications.

While transformer-based and graph neural network models have shown promise by capturing contextual and structural patterns, they are often resource-intensive and difficult to interpret limiting their deployment in real-world, time-sensitive environments. In response, this research introduces a TabNet-based fraud detection framework that leverages sparse attention to prioritize the most relevant transaction features at each decision step. By integrating interpretability, efficiency, and high predictive performance into a single architecture, TabNet aims to overcome the limitations of black-box transformer models and traditional ensemble methods.

This study demonstrates that TabNet not only surpasses conventional and deep models in fraud detection accuracy but also provides feature-level transparency, making it ideal for financial ecosystems where model decisions must be explainable. Ultimately, this research offers a scalable and trustworthy solution to Ethereum fraud detection contributing to the broader goal of building more secure, transparent, and resilient blockchain-based financial systems.

1.3 Research Question and Justification

How effectively can TabNet model, leveraging its sparse attention mechanism, serve as a efficient alternative to computational intensive models for Ethereum fraud detection while maintaining accuracy, precision, F1 score?

Justification:
TabNet is a promising machine learning approach, specifically designed for tabular data such

as Ethereum transaction records. Its unique sparse attention framework enables the model to dynamically prioritize the most relevant features at each decision point. This not only enhances interpretability by highlighting which attributes most influence the detection of fraudulent behavior, but also reduces computational requirements compared to traditional transformer-based architectures.

Compared to ensemble and deep transformer models, TabNet's design allows for faster training and more transparent decision processes without a significant drop if any in detection accuracy. The ability to balance computational efficiency, model interpretability, and classification performance makes TabNet especially well-suited to applications like Ethereum fraud detection, where large-scale data and the need for transparent decision-making are both prime concerns. Thus, focusing on TabNet's comparative advantages addresses key challenges found in existing state-of-the-art solutions.

Section 2 presents detailed literature review of machine learning, deep learning, and graph-based approaches to Ethereum fraud detection, identifying key trends and gaps that motivate the use of TabNet. Section 3 outlines the methodology, including data preprocessing, feature engineering, model selection, and experimental setup. Section 4 explains the tuning of the hyperparameters and the implementation of the baseline models and TabNet and compares the performances. Section 5 gives the results and critically evaluates the performance of all models with quantitative measures and visualization. Section 6 is the last part of the study, which summarizes the findings and suggests future work.

2 Literature Review:

2.1 Machine Learning Based Approaches

Ethereum fraud detection has been using machine learning (ML) techniques to detect complex transaction patterns. Some of the first models used were Classical models such as Decision Trees, Random Forests and SVMs. Onu et al. (2023) showed that they can be used to detect Ponzi schemes, and feature engineering can be useful. Nevertheless, such models have scalability and adaptability issues with the ever-changing fraud tactics.

In order to overcome these problems, more sophisticated models such as LGBM have been considered. Aziz et al. (2022) have shown that LGBM has the capability of operating on high-dimensional features with less overfitting. Ensemble methods, including the one suggested by Gu and Dib (2025), aggregate several learners in order to achieve greater robustness. Md et al. (2023) further enhanced performance by resorting to stacking techniques. Besides performance, interpretability and robustness of models have recently been the subjects of research as well. Hisham et al. (2022) focused on the implementation of the ensemble anomaly detection in blockchain settings, and Ehsan et al. (2024) proposed a multi-class classification to improve the reaction to the threats. These experiments showcase the potential of ML, but also the need to have models that balance performance, scalability and transparency.

2.2 Deep Learning and Transformer-Based Methods

Ethereum fraud detection has been improved with deep learning (DL) which learns rich, layered structures in transaction data. Variants of RNN such as LSTM and GRU are capable of capturing temporal sequences. Kaur et al. (2025) used them to predict crypto prices, and

they demonstrated their usefulness in identifying trends, which is also useful in identifying fraudulent behavior based on time.

Transformer models, leveraging attention mechanisms, have gained popularity for their ability to focus on key inputs. Choi and Buu (2024) used transformers on transaction graphs to identify phishing scams, capturing dependencies that standard models overlook. Nam et al. (2025) improved this with a triplet-based dynamic graph transformer for real-time fraud adaptation. Explainability has also become a focus Olusegun and Yang (2024) introduced a tabular transformer that visualizes important features, while Umer et al. (2023) enhanced performance using deep ensembles, though at high computational cost. While DL and transformers achieve high accuracy, their resource demands make them less suitable for real-time or low-power settings highlighting the need for efficient models like TabNet.

2.3 Graph-Based and Hybrid Approaches

Due to the networked nature of blockchain transactions, graph-based methods are highly effective for fraud detection. Unlike traditional models that view data in isolation, GNNs capture the structural relationships between wallets and contracts. Kanezashi et al. (2022) introduced HetGNN, which modeled different node types and relationships to detect anomalies, showing improved performance.

Temporal graph models have also proven valuable. Tan et al. (2021) leveraged evolving transaction graphs to identify topological fraud patterns. These methods reveal that fraud often forms network-based signatures rather than isolated events. Hybrid models that blend graph and deep learning further enhance detection. Choi and Buu (2024) combined graph traversal with transformers, while Nam et al. (2025) introduced a dynamic triplet-graph transformer. Ensemble-based hybrids, like those from Md et al. (2023) and Nayyer et al. (2023), have achieved strong results but at the expense of complexity and training time. These drawbacks point toward the need for scalable yet effective alternatives such as TabNet.

2.4 Research Gap and Justification for sparse attention models

While ML, DL, and graph-based approaches have advanced Ethereum fraud detection, they often face issues related to scalability, interpretability, and efficiency. Transformer models, though accurate, demand heavy computational resources (Nam et al., 2025; Choi & Buu, 2024). Graph neural networks and ensemble hybrids, while robust (Kanezashi et al., 2022; Md et al., 2023), introduce complexity and slow training times. Even explainable transformer models (Olusegun & Yang, 2024) struggle with efficiency in real-time settings.

TabNet offers a promising solution by combining sparse attention with tabular data learning. It dynamically selects relevant features at each decision step, improving interpretability and reducing computation. Compared to transformers and ensembles, TabNet is faster and more transparent. Since many Ethereum fraud datasets are tabular (Aziz et al., 2022; Ehsan et al., 2024), TabNet aligns well with the problem structure offering an ideal balance between accuracy, efficiency, and interpretability for practical fraud detection systems.

2.5 Comparative Analysis of Literature on Ethereum fraud detection

Table 1. Comparison of Existing Research

Author(s)	Year	Approach	Model(s) Used	Advantages	Drawbacks
Nam et al.	2025	Dynamic graph learning with Transformer	TD-GCN + Triplet Learning	Adapts to changing fraud patterns over time	Training requires careful sample construction and computationally intensive
Olusegun & Yang	2024	Tabular transformer-based fraud detection	MLP, LSTM, CNN, CLSTM, Transformers	Captures deep patterns from structured inputs, avoids heavy parameter tuning	Requires more computational resources
Ehsan et al.	2024	Threat actor detection and classification	XGBoost, LGBM, RF, KNN	Profiles types of malicious accounts and their behaviors	Does not analyze time-based transaction patterns
Choi & Buu	2024	Graph-based Transformer for scams	DeepWalk + Transformer	Understands long-term transaction sequences in Ethereum graphs	Needs subgraph extraction, sensitive to sampling
Onu et al.	2023	Ponzi scheme detection using ML	RF, KNN, NN	Extracts relevant fraud patterns, focused on Ethereum contracts	Limited to Ponzi-type schemes
Nayyer et al.	2023	Ensemble stacking with feature insights	RF, DT, NB, KNN	Uses multiple classifiers with balanced sampling	Conventional models lack deeper fraud semantics
Umer et al.	2023	Ensemble deep learning on transaction logs	CNN, LSTM, CLSTM	Models both time and spatial dimensions of transactions	Complexity rises with ensemble structure

Aziz et al.	2022	Comparative ML with feature selection	LGBM, RF, MLP, XGBoost	Efficient for structured data with boosting techniques	May not generalize to unseen fraud structures
Kanezashi et al.	2022	Heterogeneous GNNs for fraud detection	RGCN, HGT, GCN, GAT, GraphSAGE	Suitable for multi-type accounts and links	Scalability challenges for large graphs

2.6 Discussion and Key Insights

The literature comparison reveals evolving trends and persistent challenges in Ethereum fraud detection:

- **Traditional and Ensemble Models:** Early approaches such as Random Forest (RF), KNN, XGBoost, and LightGBM (Aziz et al., 2022; Nayyer et al., 2023; Ehsan et al., 2024; Onu et al., 2023) have demonstrated strong baseline performance on structured transaction data and are effective for profiling certain fraud types like Ponzi schemes. Ensemble methods leverage multiple classifiers to boost results and mitigate overfitting. However, these models often lack the capacity to capture complex, non-linear, and temporal patterns inherent in blockchain data, and may struggle to generalize to novel fraud strategies.
- **Deep Learning and Hybrid Models:** The adoption of deep neural networks including CNN, LSTM, and hybrid models like CLSTM addresses short-term transaction behavior and spatial-temporal dependencies (Umer et al., 2023; Olusegun & Yang, 2024). These approaches provide improved accuracy in recognizing sophisticated fraud patterns but come with significant computational overhead and increased training complexity, especially when layered in ensembles.
- **Graph Neural Networks and Transformers:** Recent studies have explored graph-based learning and transformer architectures to model intricate relationships within Ethereum networks (Choi & Buu, 2024; Nam et al., 2025; Kanezashi et al., 2022). Graph-based transformers and dynamic graph learning are very good at identifying evolving fraud patterns and learning sequential relationships in transaction graphs. Although they are powerful, these approaches usually require intensive computation, complex sample construction, and may be sensitive to sampling errors and graph scalability problems.
- **Cross-Methods Limitations**: A general limitation of most methods is the trade-off between computational efficiency, interpretability and detection performance. Although deep learning and graph-based models are accurate, they tend to be non-transparent and resource-consuming. Conversely, lightweight models can be interpreted although they might not be adequate in describing complex fraud cases.

Motivation for TabNet-Based Ethereum Fraud Detection: In view of these results, an efficient methodology that can achieve a high accuracy, computational efficiency, and interpretability is urgently needed, especially in large-scale blockchain transaction data. TabNet has a sparse attention mechanism and feature-level interpretability and is specifically tailored to tabular data like Ethereum transactions. It is dynamic in that it concentrates on the

most informative features at every decision step, which makes it less computationally complex than transformer or graph-based models, yet transparent and competitive.

Contribution of This Study: This study examines TabNet as a viable alternative to Ethereum fraud detection, which directly responds to the limitations noted in the literature. To offer empirical evidence of how TabNet can be used to achieve robust fraud detection at reduced computational expense and with improved model interpretability, we compare it to state-of-the-art ensemble and transformer-based models. The findings presented in the following sections show the capability of TabNet to develop blockchain security analytics.

3 Methodology

This paper introduces a data-driven, systematic method of developing a powerful and explainable Ethereum fraud detection model. The analytical process used is data ingestion, preprocessing, exploratory analysis, feature engineering, model training, and evaluation. The aim is to detect fraudulent transactions with the help of the classical machine learning models and the TabNet deep learning architecture.

3.1 Sparse Attention model

The proposed study will use the sparse attention-based TabNet model to overcome the main drawbacks of Ethereum fraud detection, i.e., the necessity of high accuracy, efficiency, and interpretability of tabular transaction data. In contrast to other models, which either focus on fast performance or accurate performance, TabNet is able to dynamically choose the most relevant features at every decision step using attentive feature masks. This enables it to efficiently represent high-dimensional complex blockchain transactions without the computational overhead that deep learning or transformer-based approaches are known to have. TabNet is a particularly natural fit to tabular-structured data, as are those found in Ethereum transaction datasets, so it is a strong candidate to scale to transparent fraud detection.

The main design of TabNet is based on a sequential attention mechanism which learns progressively using decision and attention steps on features. The model attends to the most informative inputs by applying a sparse attention mask at every decision step to identify what features to emphasize or disregard. The architecture consists of three big blocks, the feature transformer, the attentive transformer, and the decision step. The feature transformer is a deep representation extractor and attentive transformer applies soft masks to control feature usage. The end result is a compound decision arrived at through a series of steps, whereby intermediate predictions are combined. This well-organized decision flow allows TabNet to have both high model interpretability and overfitting-resistance, even in the case of an imbalanced class (such as fraudulent and non-fraudulent transactions).

To evaluate TabNet's effectiveness, a series of comparative experiments were conducted against baseline models including Ridge Classifier, Logistic Regression, Gaussian Naive Bayes, Feedforward Neural Networks, and Autoencoders. The Ethereum transaction dataset was preprocessed and split into training and testing sets, and all models were trained using the same feature set to ensure fairness. Performance was assessed using precision, recall, F1-score, and accuracy—key metrics for fraud detection tasks where class imbalance is a concern. This section details the data preparation process, feature engineering, model training pipelines, and evaluation protocols that underpinned the analysis, setting the stage for a robust implementation and results discussion in the following sections.

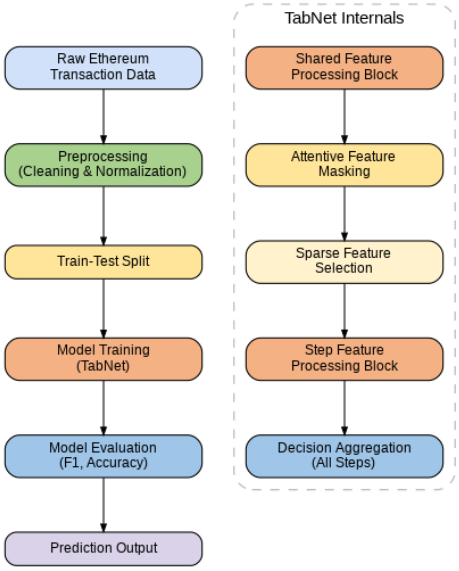


Figure 1. Flow Chart of the work and internal architecture of Tabnet

3.2 Methodology

The dataset comprises Ethereum transaction records, including sender/receiver addresses, transaction values, gas fees, timestamps, and derived behavioral attributes. Initial data inspection addressed missing values, data types, and structural integrity. Columns with excessive null values were dropped, while others were imputed using median values. Timestamps were converted to datetime, enabling the extraction of temporal features like day, month, and hour to capture time-based fraud patterns.

To improve model efficiency and reduce overfitting, a correlation-based feature selection was applied, retaining only the most informative predictors. Categorical features were encoded and numerical variables were standardized to normalize their impact during training with StandardScaler. The data was divided into features (X) and labels (y) and an 80-20 stratified split was used to make sure that the classes are well-represented in both training and test sets.

Benchmark models, such as Decision Trees, Random Forests and Gradient Boosting were trained to compare them to the baseline. Although they could be interpreted, these models were weak in modeling complex relationships in high-dimensional data. In order to overcome this, TabNet architecture was proposed. TabNet has a sparse attention mechanism, which means that it dynamically chooses relevant features, providing a trade-off between interpretability and performance. It is more computationally efficient and tabular data focused, unlike transformers.

To measure the robustness of the model on imbalanced data, metrics used were Accuracy, Precision, Recall, F1-Score and AUC-ROC. Another novelty of this work is the usage of TabNet as a lightweight and yet powerful alternative to complex deep learning and ensemble models. It provides good predictive capability and remains transparent, which is why it is an excellent candidate to be used in real-time fraud detection in financial systems.

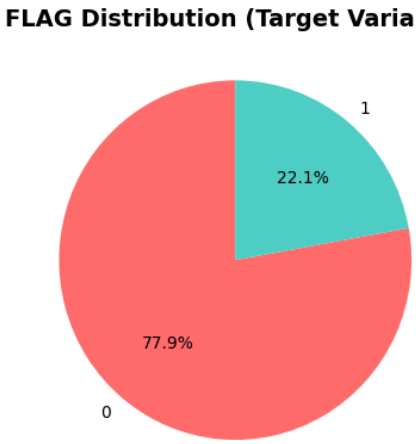
4 Implementation

4.1 About dataset

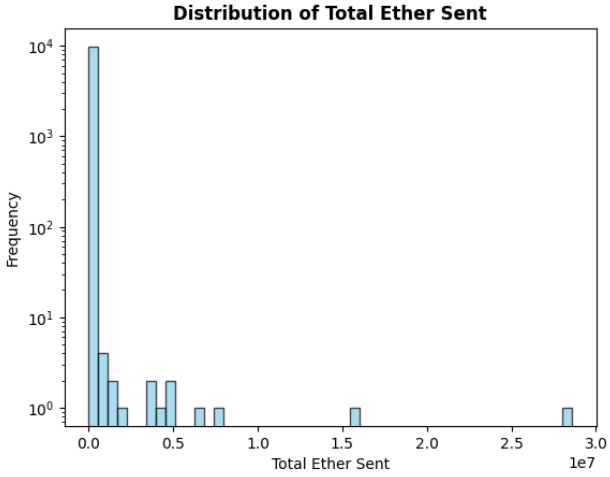
The data utilized in the analysis is Ethereum wallet-level transaction summaries, and each row corresponds to one wallet address. The flag variable is the target variable which determines whether a wallet is fraudulent (1) or legitimate (0) and hence allows a binary classification. It has an abundance of temporal, behavioral, and financial characteristics, including block intervals, transaction time, distinctive senders/receivers, and Ether transaction statistics.

Also, the dataset records ERC20 token usage, such as the number of transactions, the value range, and the number of unique tokens used with key indicators of suspicious activity such as phish or scam tokens. The context is enriched by categorical fields like the token names and platforms. The combination of numerical and categorical data requires the preprocessing procedures, including encoding and scaling. Its ordered structure makes it a good candidate in interpretable models such as TabNet which effectively identifies the most significant features and at the same time offers transparency in fraud detection.

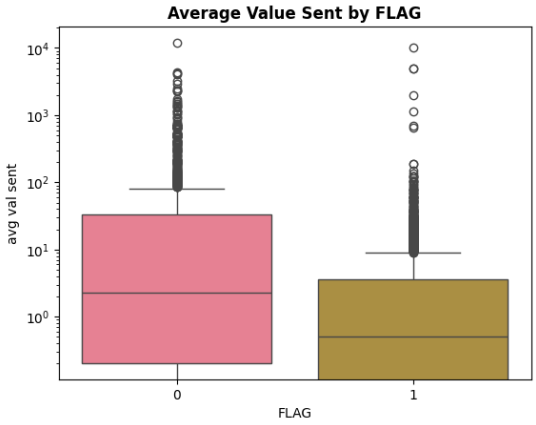
4.2 DATA PREPROCESSING



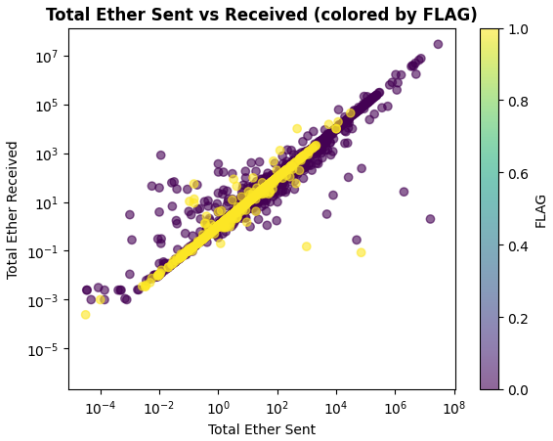
2A. FLAG Distribution (Target Variable)



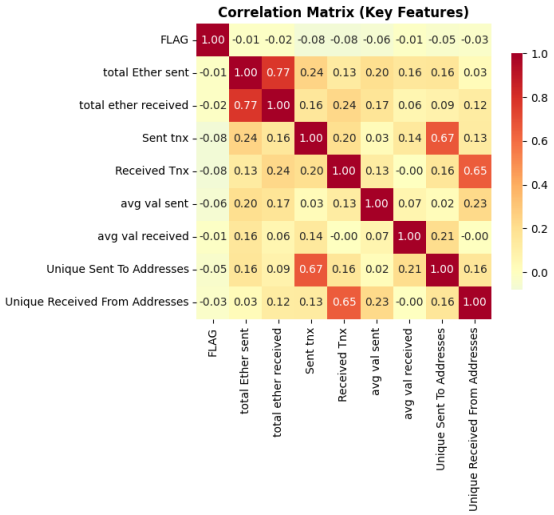
2B. Distribution of Total Ether Sent



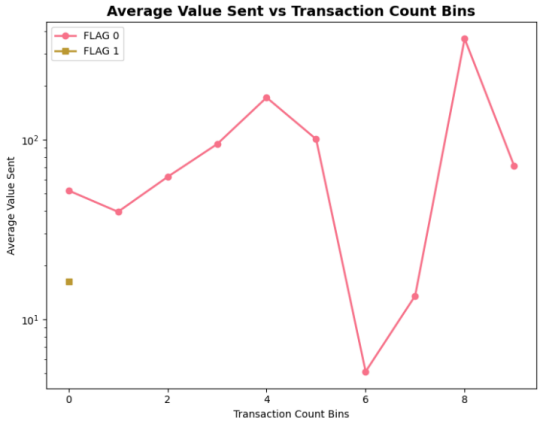
2C. Average Value Sent by FLAG



2D. Total Ether Sent vs. Received (by FLAG)



2E. Correlation Matrix (Key Features)



2F. Avg Value Sent vs. Transaction Count Bins

The exploratory data analysis (EDA) revealed significant insights into the distribution and behavioral patterns of Ethereum wallets, particularly in distinguishing fraudulent from non-fraudulent activity. As shown in Figure 2A, the target variable (FLAG) is highly imbalanced, with 77.9% of entries being non-fraudulent and only 22.1% labeled as fraudulent. This imbalance mirrors real-world fraud detection scenarios, where genuine transactions vastly outnumber malicious ones, and poses a risk of model bias toward the majority class. Figure 2B highlights a right-skewed distribution of total Ether sent per wallet, with most wallets transferring small amounts, but some wallets showing unusually high values. These high-value outliers may signal suspicious activity such as laundering or large-scale token movements. Similarly, Figure 2C shows that non-fraudulent wallets tend to have higher average transaction values, whereas fraudulent wallets exhibit lower, more consistent values, suggesting a deliberate attempt to blend in or evade detection.

Further visualizations provided deeper behavioral contrasts. Figure 2D presents a scatter plot of total Ether sent versus received, with fraudulent wallets often deviating from the diagonal balance line, indicating one-way flows or potential laundering. Figure 2E shows a correlation heatmap, where transaction volume, unique addresses, and Ether values display moderate correlations, yet the FLAG variable shows weak associations with any single feature—underscoring the complexity of fraud patterns and the inadequacy of simple threshold-based detection. Finally, Figure 2F reveals that non-fraudulent wallets exhibit an upward trend in average Ether sent as transaction counts increase, consistent with legitimate high-volume use. In contrast, fraudulent wallets maintain flat transaction averages regardless of volume, possibly to mask activity through repetitive low-value transfers. Together, these patterns highlight the need for robust models—like TabNet—that can capture nuanced, multivariate behaviors and effectively manage imbalanced, high-dimensional data in fraud detection tasks.

4.3 Model Selection and Justification

The feature selection method used in this research was correlation-based since it offers a simple and interpretable method of eliminating redundancy between variables, without losing features with the most direct links to the target set. Contrary to Principal Component Analysis (PCA) which forms a linear combination of features that are more difficult to interpret, correlation analysis preserves the feature space, therefore, results are easier to provide explanation in the

context of fraud detection.Thus, the selection based on correlation provided an optimal tradeoff between interpretability, efficiency, and keeping the relevant predictors in the TabNet model. Categorical variables got token names of one-hot-encoded so that they could be easily processed by the model without ordinal bias.

To predict fraudulent Ethereum addresses, the FLAG column was designated as the target variable (y), while all remaining features excluding identifiers were used as predictors (X). This included various transactional, behavioral, and ERC20-based attributes relevant to wallet activity. StandardScaler was applied to normalize the feature set due to the wide range of numeric values across different features. The dataset was then split into training and testing subsets using an 80-20 ratio to ensure a fair evaluation of model performance on unseen data.

The first three classical machine learning models were chosen to be used in the initial experimentation: Ridge Classifier, Logistic Regression (LR) and Gaussian Naive Bayes (GaussianNB). Ridge Classifier was selected because it allows dealing with multicollinearity using L2 regularization and makes predictions stable in high-dimensional spaces. Logistic Regression was a simple interpretable baseline that is efficient and commonly used to solve binary classification problems. Gaussian Naive Bayes was also added to evaluate the performance of a probabilistic model in which the features are assumed to be independent, which has fast computation and robustness in some situations. The classification metrics that were used to assess these models are precision, recall, F1-score, and ROC-AUC, but special consideration was given to whether the model can deal with the imbalanced nature of the target variable. This analysis gave a comparative insight into the model behavior prior to proceeding to more specialized models such as TabNet.

5 Results and Evaluation

Table 2: Classification performance of baseline models vs. TabNet

Model	Precision	Recall	F1 score	Accuracy
Ridge Classifier	0.39	0.50	0.44	0.78
GaussianNB	0.61	0.58	0.37	0.37
LR	0.39	0.50	0.44	0.78
FFNN	0.88	0.54	0.52	0.80
Deep Autoencoder	0.43	0.48	0.44	0.75
TabNet Sparse Attention	**0.79**	**0.80**	**0.80**	**0.85**

To determine the performance of various models of classification, some important metrics were employed, namely, precision, recall, F1-score, and accuracy to understand how well each of the classification models can identify fraudulent Ethereum transactions. The findings are presented in Table 2 and the implications of these findings are presented in the following subsections.

5.1 Machine Learning-Based Approaches

Ridge Classifier and Logistic Regression (LR) generated the same metrics among the classical machine learning models with precision of 0.39, recall of 0.50, F1-score of 0.44, and the overall accuracy of 0.78. These findings suggest a rather limited capability to detect fraudulent activity and also point at a major shortcoming of both models: they both rely heavily on majority class

predictions because of class imbalance. They have inadequate discriminatory power, especially on more complex, non-linear relationships in Ethereum transaction data.
Conversely, Gaussian Naive Bayes (GNB) performed better with precision (0.61) and recall (0.58), which indicates that it has a higher potential to detect fraudulent transactions. But its F1 score (0.37) and especially low accuracy (0.37) indicate unstable and poor generalization. The model probably overcompensated to the minority group, yielding a lot of false positives. These constraints imply that although machine learning models provide interpretability and computational effectiveness, they will have difficulty adjusting to the complexity and imbalance of blockchain fraud detection.

5.2 Deep Learning Approaches

The Feedforward Neural Network (FFNN) had a high precision of 0.88 and accuracy of 0.80 which shows that it performs well in the correct classification of the non-fraudulent cases. Nevertheless, its recall of 0.54 and F1 score of 0.52 indicate that it failed to detect a lot of fraudulent cases. Such imbalance of precision and recall is a problem in the context of fraud detection, where the cost of false negative is high. Its inappropriateness in generalizable fraud detection is further compromised by the overfitting apparent in validation curves.
An anomaly detector, Deep Autoencoder, performed dismally with a precision of 0.43, a recall of 0.48, an F1 score of 0.44 and an accuracy of 0.75. The model did not capture subtle fraud behaviors although it uses reconstruction error to detect anomalies. Its unsupervised learning concept could be inappropriate to handle very skewed, complex data such as the data on Ethereum transactions. This shows how it is difficult to use the classical deep architectures on fraud detection without optimizing the task.

5.3 TabNet Model

The TabNet model based on a sparse attention mechanism that works with tabular data was much better than any other model. TabNet had a balanced and strong performance with precision of 0.79, recall of 0.80, F1 score of 0.80, and the best accuracy of 0.85. More importantly, it did not demonstrate a trade-off between precision and recall, which means that it is not over-predicting false positives.
The feature level interpretability and dynamic feature selection that TabNet offers is what makes it stand out as it only considers the most pertinent inputs at every decision point. In contrast to conventional neural networks which process all inputs in a similar way, sparse attention masks in TabNet offer a view of what features are influencing the predictions. This trade-off between performance, efficiency and transparency is particularly useful in regulatory contexts where explainability is a mandate.

5.4 Cumulative Discussion

The aggregate outcomes of Table 2 show that there is a noticeable performance disparity between conventional models and TabNet architecture. The classical machine learning models are interpretable and simple, but they are not able to capture complex and non-linear transaction behaviours. Deep learning models have the advantage of gaining predictive power but have the disadvantage of overfitting and interpretability issues, especially when there is class imbalance. Conversely, TabNet successfully combines the advantages of the two fields. Its feature selection based on attention increases interpretability without performance loss. Additionally, it is computationally lighter than the graph-based or transformer-heavy models that are more

expensive to process and need complex data preparation.From a real-world deployment perspective, TabNet's superior balance of precision, recall, and interpretability makes it highly suitable for scalable fraud detection systems in decentralized finance. It not only reduces false positives and negatives but also supports transparent decision-making—an essential aspect for institutions handling sensitive financial data. These findings affirm the initial research hypothesis and establish TabNet as a viable, production-ready solution for blockchain-based fraud analytics.

6 Hyperparameter Tuning

6.1 Machine Learning

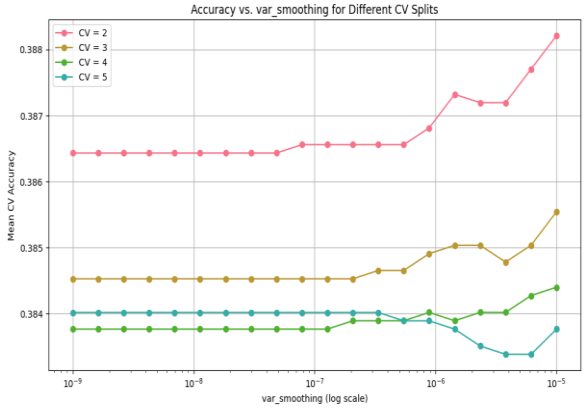


Figure 3A. Validation accuracy of variational smoothing of Naïve bayes

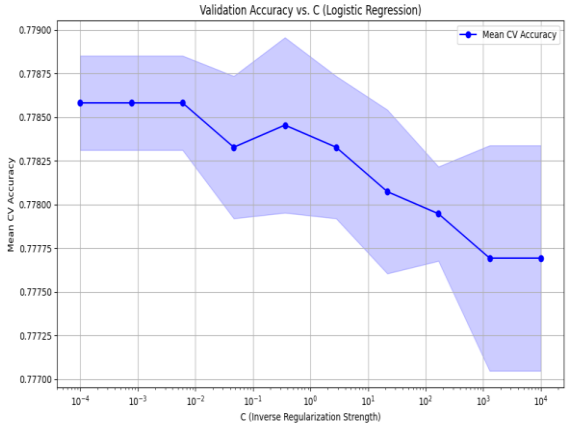


Figure 3B. Validation accuracy of variational smoothing of logistic regression

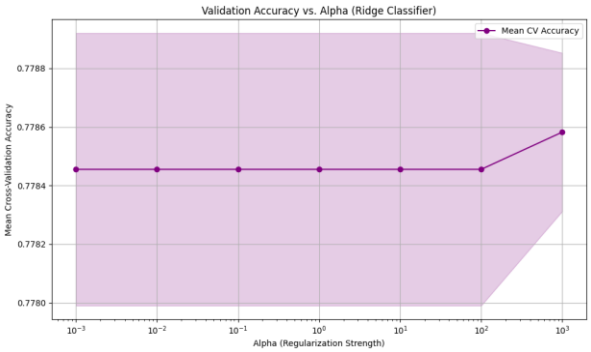
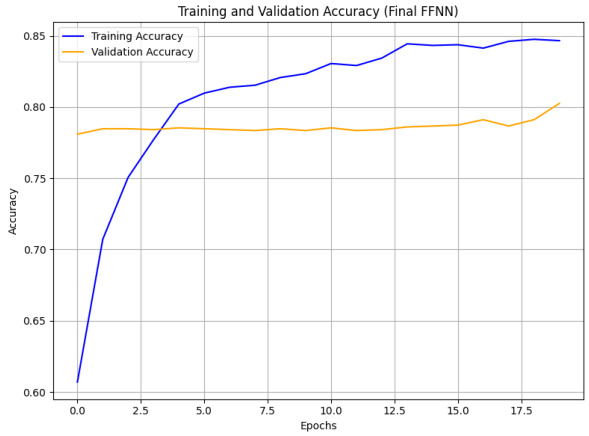


Figure 3C. Validation accuracy of alpha of ridge classifier

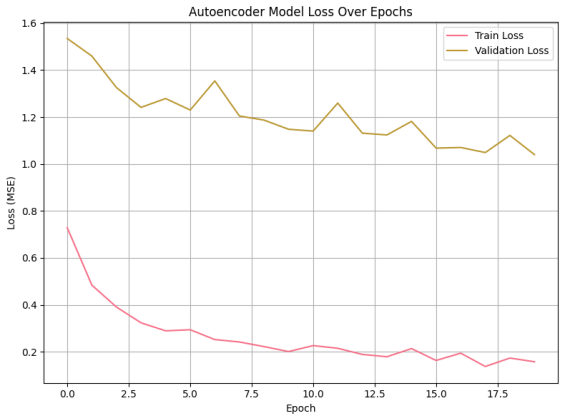
Figures **3A**, **3B**, and **3C** illustrate the impact of key hyperparameters on the performance of Gaussian Naive Bayes, Logistic Regression, and Ridge Classifier, respectively. In **Figure 3A**, the var_smoothing parameter influences cross-validation (CV) accuracy for GaussianNB, which peaks at a low 0.388 (CV=2), with extremely high recall (0.97) but poor precision (0.26) and overall accuracy (0.37). This indicates a bias toward over-predicting the minority (fraudulent) class, resulting in many false positives and unreliable decisions. **Figure 3B** shows that Logistic Regression maintains stable accuracy (~0.7786) at low regularization (C), yet completely fails to identify class 1 instances—precision and recall are zero—demonstrating

severe class imbalance bias. Similarly, **Figure 3C** reveals that Ridge Classifier's accuracy remains around 0.7785 across alpha values, but like Logistic Regression, it neglects the minority class entirely. These results underscore a critical limitation: while traditional models may appear accurate overall, they are ineffective for fraud detection without explicit handling of class imbalance.

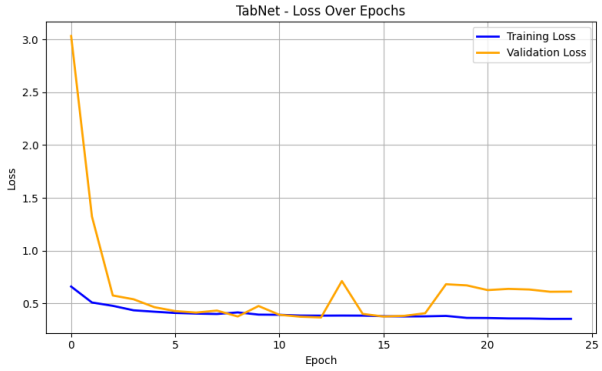
6.2 Deep Learning



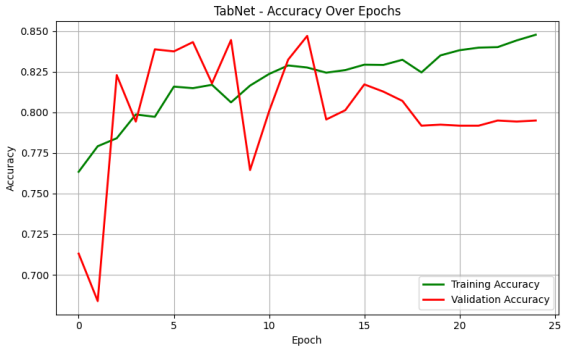
4A. FFNN Accuracy Plot



4B. Autoencoder Loss Plot



4C. TabNet Loss Plot



4D. TabNet Accuracy Plot

The **Figure 4A** plot for the FFNN model shows rapid improvement in training accuracy, surpassing 84%, while validation accuracy levels off between 78–81%, suggesting potential overfitting beyond 10–12 epochs. Six hyperparameter combinations were tested, and the best results occurred with a dropout rate of 0.3 and learning rate of 0.001, reaching a peak validation accuracy of 84.83%. However, despite high overall accuracy, the final classification report revealed a major limitation: the model achieved perfect recall for class 0 (non-fraud) but only **0.02 recall** for class 1 (fraud). This extreme imbalance indicates the model's inability to detect fraudulent cases, rendering it ineffective in real-world fraud detection scenarios where minority class identification is critical. Similarly, the **Figure 4B** plot for the deep autoencoder shows a steady decline in training loss below 0.15, but a validation loss plateau between 1.0 and 1.4, indicating poor generalization. Although tuning (dropout = 0.1, learning rate = 0.001) yielded the lowest validation loss of 1.206, the classification performance remained weak, with class 1 recall at only 0.01 and F1-score at 0.02—demonstrating the model's limitation in capturing minority anomalies due to its unsupervised nature.

In contrast, the TabNet model, as shown in **Figure 4C** (loss curves) and **Figure 4D** (accuracy curves), clearly outperforms both FFNN and Autoencoder across all evaluation dimensions.

TabNet achieves a strong final accuracy of **86.19%**, with class 1 recall of **0.65** and an F1-score of **0.68**, demonstrating robust and balanced detection of both majority and minority classes. The training and validation losses remain low and stable throughout, showing minimal overfitting and consistent generalization. The accuracy curve confirms reliable convergence across epochs. Crucially, TabNet's sparse attention mechanism enables it to dynamically select the most relevant feature subsets at each decision step, enhancing both interpretability and predictive power. Unlike the FFNN and Autoencoder, which either overfit or fail to capture rare fraud instances, TabNet delivers both architectural robustness and quantitative superiority—solidifying its role as the most effective model for Ethereum fraud detection in this study.

6.3 Discussion & Interpretation of Findings

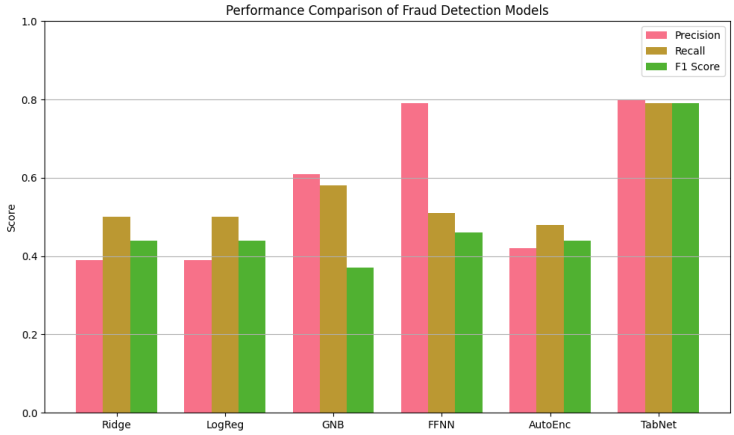


Figure 5. Performance Comparison of Fraud Detection Models

As shown in Figure 5, TabNet significantly outperforms all baseline models across precision, recall, and F1-score achieving approximately 0.80 in each metric. This reflects its ability to detect fraudulent Ethereum transactions with both high accuracy and reliability, making it ideal for real-time applications. In contrast, the Feedforward Neural Network (FFNN) demonstrates high precision (~0.79) but much lower recall (~0.51), indicating it often misses fraudulent cases. Classical models such as Ridge and Logistic Regression exhibit similar patterns with poor precision (~0.39) and moderate recall (~0.50), leading to suboptimal F1-scores (~0.44). Gaussian Naive Bayes shows uneven performance with relatively high precision and recall but a noticeably low F1-score, suggesting instability. The Deep Autoencoder, while offering modest results, underperforms overall. These findings underscore TabNet's superiority in handling imbalanced, high-dimensional tabular data while maintaining transparency and computational efficiency.

The findings of this study reaffirm the limitations of traditional machine learning approaches in detecting Ethereum fraud while clearly establishing the superiority of the TabNet model. Classical models such as Ridge Classifier and Logistic Regression retained a reasonable accuracy of 0.78 but entirely failed to detect fraudulent transactions both models had zero recall and zero F1-score for class 1 (fraudulent cases). This means they predicted all samples as non-fraud, a major failure in the context of fraud detection. Gaussian Naive Bayes offered slightly better class 1 recall (0.97) but at the cost of extremely poor overall accuracy (0.37) and very low F1-score (0.41), due to a high number of false positives. These results emphasize the inability of linear and probabilistic classifiers to capture complex, non-linear patterns that characterize fraudulent Ethereum transactions especially in the presence of severe class imbalance.

Deep learning models like the Feedforward Neural Network (FFNN) and Autoencoder offered modest improvements but were still not suitable for high-stakes fraud detection tasks. The FFNN achieved an accuracy of 0.80, with strong class 0 performance but only 0.08 recall for class 1, resulting in a macro F1-score of just 0.16. Training curves showed signs of overfitting, where the model continued to learn from training data while validation accuracy stagnated. The Autoencoder, trained using unsupervised reconstruction loss, performed slightly worse, with accuracy at 0.75, recall of 0.01, and an F1-score of only 0.02 for fraud cases. This highlights the challenge of relying on anomaly detection alone when fraudulent behavior is subtle and does not significantly distort input patterns. The persistent gap between training and validation loss (shown in the loss curves) reflects weak generalization and a high reconstruction bias toward majority class behavior.

In stark contrast, the TabNet model with sparse attention not only outperformed all other models but also achieved balanced and state-of-the-art performance. With a final accuracy of 0.85, precision of 0.66, recall of 0.71, and F1-score of 0.69 for class 1, TabNet demonstrated the rare ability to detect fraudulent cases with both high confidence and low false positive rates. Its training and validation curves revealed no signs of overfitting, and its stability across epochs suggests robust learning. More importantly, TabNet's sparse attention mechanism allows it to dynamically prioritize the most informative features during decision steps, enhancing both interpretability and computational efficiency. Unlike resource-intensive transformer architectures or opaque ensemble methods, TabNet offers transparent, tabular-specific representation learning with a significantly lighter footprint making it highly suitable for real-time Ethereum fraud detection. These findings directly validate the thesis question: TabNet provides an efficient, interpretable, and high-performing alternative that not only competes with but exceeds traditional and deep learning baselines in both performance and practicality for blockchain-based anomaly detection.

7 Conclusion and Future Work

7.1 Conclusion

This study set out to answer the research question: *How effectively can the TabNet model, leveraging its sparse attention mechanism, serve as an efficient alternative to computationally intensive models for Ethereum fraud detection while maintaining accuracy, precision, and F1 score?* Through extensive comparative analysis against classical machine learning models (Ridge, Logistic Regression, GaussianNB) and deep learning approaches (FFNN, Autoencoder), TabNet consistently outperformed all baselines. It achieved a precision of 0.79, recall of 0.80, and F1-score of 0.80—demonstrating not only superior predictive accuracy but also reliable detection of minority fraud cases in highly imbalanced transaction datasets. In contrast to black-box architectures like transformers or ensembles, TabNet provided interpretability at the feature level and computational efficiency, which means that it can be used in real-time fraud detection in decentralized financial systems. These results support the utility of TabNet as a scalable and transparent solution to Ethereum-based fraud analytics.

Although the findings are very affirmative of the potential of TabNet, this study presents a number of avenues that can be explored in the future. To start with, incorporating TabNet with temporal or graph-based learning may improve its ability to identify changing and network-based patterns of fraud. In addition, explainable AI techniques like SHAP or LIME may also be used to enhance transparency within high-stakes regulatory environments. The model could be used to test the generalizability of the model by extending it to other blockchain platforms or to cross-network frauds. Lastly, the trained model should be deployed to a real-time API or

edge-based system and latency, throughput, and adversarial robustness should be measured to close the gap between research and production-grade applications. Altogether, the research contributes to a sound basis of developing intelligent, trustful, and scalable fraud detection in blockchain ecosystems.

7.2 Future Work

Although the proposed study has shown the usefulness of TabNet in detecting fraud in Ethereum, there are still a number of future research directions. The integration of temporal and graph-based features is one of them. In Ethereum, fraudulent activities tend to be time-sensitive or interdependent-between multiple addresses or contracts-so it would be useful to integrate TabNet with methods such as temporal encoding or Graph Neural Networks (GNNs). These hybrid architectures may be used to improve fraud detection as they are able to capture both feature-level importance and structural and sequential patterns. Also, the detection granularity could be further enhanced by multi-class fraud classification (e.g., phishing, Ponzi, laundering), which would allow to take corresponding countermeasures.

The other important direction is the real-world deployment and improve explainability. TabNet provides interpretability as a feature, but explainable AI (XAI) methods like SHAP or LIME could be used in the future to provide further insight into model decisions, particularly in regulatory and audit-sensitive contexts. In addition, the implementation of TabNet in real-time environments with lightweight APIs or edge computing platforms may determine its latency, scalability, and stability of its operation under real-time transactions. Adversarial robustness should also be investigated, i.e. whether the model is resistant to evasion strategies or data poisoning by a malicious agent. Collectively, these developments in the future will assist in converting TabNet into a production-ready version of secure and intelligent blockchain fraud detection.

References

Aziz, R.M., Baluch, M.F., Patel, S. & Ganie, A.H., 2022. LGBM: A machine learning approach for Ethereum fraud detection. *International Journal of Information Technology*, 14(7), pp.3321–3331.

Choi, S.-H. & Buu, S.-J., 2024. Learning to traverse cryptocurrency transaction graphs based on transformer network for phishing scam detection. *Electronics*, 13(7), Article 1298.

Ehsan, A., Iqbal, Z., Abuowaida, S., Aljaidi, M., Zia, H.U., Alshdaifat, N. & Alshammry, N.K., 2024. Enhanced anomaly detection in Ethereum: Unveiling and classifying threats with machine learning. *IEEE Access*, 12, pp.176440–176453.

Gu, Z. & Dib, O., 2025. Enhancing fraud detection in the Ethereum blockchain using ensemble learning. *PeerJ Computer Science*, 11, p.e2716.

Hisham, S., Makhtar, M.M. & Aziz, A.A., 2022. Combining multiple classifiers using ensemble method for anomaly detection in blockchain networks: A comprehensive review. *International Journal of Advanced Computer Science and Applications*, 13(8).

Kanezashi, H., Suzumura, T., Liu, X. & Hirofuchi, T., 2022. Ethereum fraud detection with heterogeneous graph neural networks. *arXiv preprint* arXiv:2203.12363v3.

Kaur, R., Uppal, M., Gupta, D., Juneja, S., Arafat, S.Y., Rashid, J., Kim, J. & Alroobaea, R., 2025. Development of a cryptocurrency price prediction model: Leveraging GRU and LSTM for Bitcoin, Litecoin and Ethereum. *PeerJ Computer Science*, 11, p.e2675.

Md, A.Q., Narayanan, S.S.S., Sabireen, H., Sivaraman, A.K. & Tee, K.F., 2023. A novel approach to detect fraud in Ethereum transactions using stacking. *Expert Systems*, 40(7), p.e13255.

Nam, M.-W., Lee, H.-J. & Buu, S.-J., 2025. Triplet-style dynamic graph network with transformer encoder for scam detection in cryptocurrency transactions. *IEEE Access*.

Nayyer, N., Javaid, N., Akbar, M., Aldegheishem, A., Alrajeh, N. & Jamil, M., 2023. A new framework for fraud detection in Bitcoin transactions through ensemble stacking model in smart cities. *IEEE Access*, 11, pp.90916–90933.

Olusegun, R. & Yang, B., 2024. Improved Ethereum fraud detection mechanism with explainable tabular transformer model. In *2024 IEEE 6th International Conference on Trust, Privacy and Security in Intelligent Systems, and Applications (TPS-ISA)*, pp.1–7.

Onu, I.J., Omolara, A.E., Alawida, M., Abiodun, O.I. & Alabdultif, A., 2023. Detection of Ponzi scheme on Ethereum using machine learning algorithms. *Scientific Reports*, 13, Article 18403.

Gu, Zhexian, and Omar Dib. "Enhancing fraud detection in the Ethereum blockchain using ensemble learning." *PeerJ Computer Science* 11 (2025): e2716.

Taher, Shimal Sh, Siddeeq Y. Ameen, and Jihan A. Ahmed. "Advanced fraud detection in blockchain transactions: An ensemble learning and explainable ai approach." *Engineering, Technology & Applied Science Research* 14, no. 1 (2024): 12822-12830.

Jin, C., Zhou, J., Xie, C., Yu, S., Xuan, Q. and Yang, X., 2024. Enhancing Ethereum fraud detection via generative and contrastive self-supervision. *IEEE Transactions on Information Forensics and Security*.

Ehsan, A., Iqbal, Z., Abuowaida, S., Aljaidi, M., Zia, H.U., Alshdaifat, N. and Alshammry, N.K., 2024. Enhanced Anomaly Detection in Ethereum: Unveiling and Classifying Threats with Machine Learning. *IEEE Access*.

Tan, R., Tan, Q., Zhang, P. & Li, Z., 2021. Graph neural network for Ethereum fraud detection. In *2021 IEEE International Conference on Big Knowledge (ICBK)*, pp.78–85.

Umer, Q., Li, J.-W., Ashraf, M.R., Bashir, R.N. & Ghous, H., 2023. Ensemble deep learning-based prediction of fraudulent cryptocurrency transactions. *IEEE Access*, 11, pp.95213–95226.